OUTPUT SLACKS-ADJUSTED COST EFFICIENCY AND VALUE-BASED TECHNICAL EFFICIENCY IN DEA MODELS

Hirofumi Fukuyama, William L. Weber
Fukuoka University, Southeast Missouri State University

(Received March 30, 2008; Revised August 6, 2008)

Abstract This paper addresses two potential problems that arise when measuring Farrell cost efficiency using linear programming methods. First, when an optimal solution to the cost function allows slack in the output constraints that define the technology it is possible to increase at least one output without increasing costs. Therefore, two firms might be deemed equally cost efficient even though one firm produces more of at least one output. Second, previous research has shown that when input prices vary across firms, it is possible for firms with higher costs to be deemed more cost efficient than firms who have lower costs. We present a new measure of cost efficiency that accounts for both of these potential problems. We illustrate our method using data on Japanese securities firms operating in the 2004 to 2006.

Keywords: DEA, output slacks-adjusted cost efficiency, Japanese securities firms.

1. Introduction

Data envelopment analysis (DEA) is a mathematical programming technique for assessing the efficiency of DMUs (decision making units) that use multiple inputs to produce multiple outputs. Charnes, Cooper and Rhodes (CCR [4]) and Banker, Charnes and Cooper (BCC [2]) coined the term DEA in their extension of Farrell’s [11] technical efficiency measurement technique. Since the appearance of the CCR and BCC papers, there have been numerous articles that have applied and extended DEA. The attached CD-ROM in Cooper, Seiford and Tone [5] provide comprehensive bibliographies of research papers using DEA.

Farrell cost efficiency equals the ratio of minimum cost to actual cost. In the problem solving minimum costs, DMUs are assumed to choose input quantities given input prices and output quantities. However, Tone [24] showed that if two DMUs face different input prices but produce the same output, a DMU with higher costs can be deemed more Farrell efficient than a DMU with lower costs. Instead of choosing physical input quantities, Tone proposed a cost efficiency measure where DMUs minimize costs by choosing the amount to spend on each input, rather than by choosing physical input quantities. Tone’s method guarantees that firms with higher costs are always less efficient.

However, Tone’s method ignores the possibility of slack in the output constraints that define the DEA technology. The problem of output slack arises when the solution to the cost minimization problem involves a non-binding output constraint. Output slack exists when at least one output can be expanded at zero cost, resulting in a biased estimate of cost efficiency. The problem of accounting for slack has been incorporated in measures of technical efficiency by Tone [23], Färe and Lovell [8], and Pastor, Ruiz, and Sirvent [21]. In this paper, we extend these researchers’ work and construct new measures of cost efficiency and technical efficiency that account for the problem of output slack and also incorporate...
the critique of Tone [24] regarding Farrell cost efficiency. Our measure of cost efficiency equals the product of a radial index of minimum costs to actual costs and a non-radial index of the bias in cost efficiency that arises from ignoring output slack. We provide a similar decomposition for our technical efficiency measure.

Technical efficiency measures that incorporate slack have been used in recent financial institution efficiency studies. Such studies include Tone and Sahoo [25] for Indian life insurance companies, Liu and Tone [19] for Japanese banks, Drake, Hall and Simper [7] for Hong Kong banks, and Avkiran [1] for Australian and New Zealand banks. We estimate our new measures of cost efficiency and technical efficiency using data on Japanese securities firms operating during fiscal years 2004 to 2006. Our findings indicate that for Japanese securities firms, cost efficiency indexes that ignore output slack significantly overstate performance. We also find that the managers of securities firms systematically overspend on labor relative to capital, a finding that is consistent with Williamson’s [27] expense preference theory.

2. Framework

2.1. Standard (Farrell) cost efficiency

Let $x_{nj}$ ($n = 1, \ldots, N$) and $y_{mj}$ ($m = 1, \ldots, M$) correspond to the observed $N$ input quantities and $M$ output quantities which are assumed to be positive for all $j = 1, \ldots, o, \ldots, J$ DMUs.

The vectors $x_{no}$ ($n = 1, \ldots, N$) and $y_{mo}$ ($m = 1, \ldots, M$) represent the input and output quantities for DMU $o$ whose efficiency is to be evaluated. While inputs and outputs can be used in varying amounts consistent with the technology, we control for the presence of an uncontrollable (quasi-fixed) input that is part of the production process. Let $q \in \mathbb{R}^+$ represent this uncontrollable input. We assume strictly positive input and output quantities.

The variable returns to scale technology is represented by

$$T_{xy}(q) = \left\{ (x, y) = (x_1, \ldots, x_N, y_1, \ldots, y_M) : \sum_{j=1}^{J} x_{nj} \lambda_j \leq x_n, n = 1, \ldots, N; \sum_{j=1}^{J} q_j \lambda_j \leq q; \right.$$
$$\sum_{j=1}^{J} y_{mj} \lambda_j \geq y_m, m = 1, \ldots, M; \sum_{j=1}^{J} \lambda_j = 1; \lambda_j \geq 0, j = 1, \ldots, J \right\}$$

(1)

where the $\lambda_j$’s are intensity variables that form linear combinations of observed inputs and outputs, which, when summed to one, allow variable returns to scale. Relative to (1), the standard (Farrell) DEA cost efficiency can be quantified by

$$\gamma_{Farr} = \text{minimize}_{x,\lambda} \frac{\sum_{n=1}^{N} w_{no} x_n}{\alpha_o}$$
$$\text{subject to} \sum_{j=1}^{J} x_{nj} \lambda_j \leq x_n, n = 1, \ldots, N; \sum_{j=1}^{J} q_j \lambda_j \leq q_o;$$
$$\sum_{j=1}^{J} y_{mj} \lambda_j \geq y_m, m = 1, \ldots, M;$$
$$\sum_{j=1}^{J} \lambda_j = 1; \lambda_j \geq 0, j = 1, \ldots, J; x_n \geq 0, n = 1, \ldots, N.$$

(2)

where $w_{no}$ is the $n$-th input price that DMU $o$ faces and $\alpha_o = \sum_{n=1}^{N} w_{no} x_{no}$ is the actual cost.
2.2. Cost allocation efficiency and output slacks-adjusted cost efficiency

When input prices are not observed by the researcher, the standard framework given in (2) is inappropriate. Accounting for this possibility, Tone [24], Fukuyama and Weber [13, 14] and Tone and Sahoo [25] used the input cost-based technology. This technology set corresponds to the alternative combinations of outputs that can be produced from amounts spent on inputs given the fixed input and is written as

\[ T(q) = \left\{ (c, y) = (c_1, \ldots, c_N, y_1, \ldots, y_M) : \sum_{j=1}^{J} c_{nj} \lambda_j \leq c_n, n = 1, \ldots, N; \sum_{j=1}^{J} q_{j} \lambda_j \leq q; \right\} \]

where \( c_n \) is the input cost (expenditure) of \( n \)-th input and \( c \) is an \( N \)-dimensional input cost vector. Equivalent to (3) are the input cost set, \( V(y; q) = \{ c : (c, y) \in T(q) \} \), and the input cost-based output possibility set, \( P(c; q) = \{ y : (c, y) \in T(q) \} \). Given these representations we have \((c, y) \in T(q) \iff c \in V(y; q) \iff y \in P(c; q)\). Note that the constraint \( \sum_{j=1}^{J} \lambda_j = 1 \) in (3) models variable returns to scale for an input cost-based technology.

The input cost-based model was first employed by Färe and Grosskopf [9]. Recently, Tone [24] suggested a cost allocation efficiency framework in which the production technology is represented by (3), since standard (Farrell) cost efficiency sometimes shows that firms with higher costs are more cost efficient.

Relative to (3), the cost allocation efficiency index for DMU \( o \) is computed as

\[ \gamma_{CA}^{o} = \min_{c, \lambda} \frac{\sum_{n=1}^{N} c_n}{\alpha_{o}} \]

subject to

\[ \sum_{j=1}^{J} c_{nj} \lambda_j \leq c_n, n = 1, \ldots, N; \sum_{j=1}^{J} q_{j} \lambda_j \leq q_{o}; \]

\[ \sum_{j=1}^{J} y_{mj} \lambda_j \geq y_{m}, m = 1, \ldots, M; \]

\[ \sum_{j=1}^{J} \lambda_j = 1; \lambda_j \geq 0, j = 1, \ldots, J; c_n \geq 0, n = 1, \ldots, N. \]

where actual total costs are represented by \( \alpha_{o} = \sum_{n=1}^{N} c_{no} \). Färe and Grosskopf [9] showed that Farrell cost efficiency in (2) and the cost allocation efficiency index in (4) are the same if every DMU faces the same input prices.

When estimating (4), it is possible to encounter nonzero slacks of outputs. Output slack exists when it is possible to increase at least one output at no additional cost and occurs when the solution to (4) has at least one non-binding output constraint: \( \sum_{j=1}^{J} y_{mj} \lambda_j > y_{mo} \). In such cases, the efficiency score is biased in that two DMUs might be deemed equally cost efficient, but one DMU produces more of at least one output and no less of all other outputs than the other DMU.

For DMU \( o \), the input cost vector that minimizes total costs is no more expensive than actual costs. This condition means the objective of (4) satisfies \( \sum_{n=1}^{N} c_n/\alpha_{o} \leq 1 \), where
\(\frac{c_n}{\alpha_o}\) represents cost deflated expenditure on input \(n\). To obtain output slacks-adjusted cost efficiency divide the objective of (4) by an index of output slacks to obtain:

\[
\gamma_{o}^{OSAC} = \minimize_{c,s^+,\lambda} \frac{\sum_{n=1}^{N} c_n/\alpha_o}{1 + \frac{1}{M} \sum_{m=1}^{M} s^+_m}
\]

subject to

\[
\begin{align*}
\sum_{j=1}^{J} c_{nj} \lambda_j &\leq c_n, n = 1, \ldots, N; \\
\sum_{j=1}^{J} q_{j} \lambda_j &\leq q_o; \\
\sum_{j=1}^{J} y_{mj} \lambda_j &\leq y_{mo} + s^+_m, m = 1, \ldots, M; \\
\sum_{j=1}^{J} \lambda_j &= 1; \lambda_j \geq 0, j = 1, \ldots, J; \\
c_n &\geq 0, n = 1, \ldots, N; s^+_m \geq 0, m = 1, \ldots, M.
\end{align*}
\]

where the variables \(s^+_m\) represent output slacks (underproduction of outputs). The value of \(\gamma_{o}^{OSAC}\) includes the size of output slacks. In contrast, the cost allocation efficiency measure \(\gamma_{o}^{CA}\) does not account for output slacks. When there is no slack in the output constraints, the cost efficiency measures (5) and (4) are equivalent. If \(\gamma_{o}^{OSAC}\) and \(\gamma_{o}^{CA}\) are feasible and bounded, then the relation between the two cost efficiency measures is

\[
\gamma_{o}^{OSAC} \leq \gamma_{o}^{CA} \leq 1
\]

for \((c_o, y_o) \in T(q)\). Using (6), we define the bias in measured cost efficiency arising from output slacks as

\[
B_{o}^{CE} = \frac{\gamma_{o}^{OSAC}}{\gamma_{o}^{CA}}.
\]

The bias index shows the overestimate of cost allocation efficiency that arises from ignoring output slacks. Higher values of the index indicate less bias. Rearranging (7), output slacks-adjusted cost efficiency\(^*\) is decomposed into the product of cost efficiency bias and cost allocation efficiency:

\[
\gamma_{o}^{OSAC} = B_{o}^{CE} \times \gamma_{o}^{CA}.
\]

Tone [23] presents a slacks-based measure of technical efficiency that utilizes input quantities needed to produce give output quantities. We modify his slacks-based technical efficiency measure to a technical efficiency measure based on the input values (expenditures) needed to produce given physical output quantities. We refer to this index as value-based technical efficiency (VTE). The linear programming problem that estimates this index takes the form

\[
\rho_{o}^{VTE} = \minimize_{s^-,\lambda, s^+} \frac{1 - \frac{1}{N} \sum_{n=1}^{N} s^-_n}{1 + \frac{1}{M} \sum_{m=1}^{M} s^+_m}
\]

\(^*\)The standard cost efficiency decomposition into technical and allocative efficiency was proposed by Kopp and Dievert [18]. Zieschang [29] established the relation between the standard cost efficiency decomposition and the duality between the cost function and the technology.
subject to \[ \sum_{j=1}^{J} c_{nj} \lambda_j = c_{no} - s_n^-, n = 1, \ldots, N; \sum_{j=1}^{J} q_j \lambda_j \leq q_o; \] \[ \lambda_j \geq 0, j = 1, \ldots, J; s_n^- \geq 0, n = 1, \ldots, N; s_m^+ \geq 0, m = 1, \ldots, M. \]

where \( s_n^- = c_{no} - \sum_{j=1}^{J} c_{nj} \lambda_j \) are the input cost slacks which represent the additional cost due to the excessive use of at least one input. We refer to (9) as value-based technical efficiency because the physical input vector is replaced by the input cost vector. The denominator of (5) is the same as the denominator of (9).

Equivalent to (9) is a value-based version of Pastor, Ruiz and Sirvent’s [21] enhanced Russell (ERuss) technical efficiency:

\[ \rho_o^{ERuss} = \min_{\theta, \phi, \lambda} \frac{1}{N} \sum_{n=1}^{N} \theta_n \]

subject to \[ \sum_{j=1}^{J} c_{nj} \lambda_j \leq \theta_n c_{no}, n = 1, \ldots, N; \sum_{j=1}^{J} q_j \lambda_j \leq q_o; \] \[ \phi_m \geq 1, m = 1, \ldots, M. \]

where the choice variables, \( \theta = (\theta_1, \ldots, \theta_N) \) and \( \phi = (\phi_1, \ldots, \phi_M) \) are non-symmetric scaling factors for input costs and outputs.

Define value-based Farrell technical efficiency as

\[ \rho_o^{Far} = \min_{\rho, \lambda} \rho \]

subject to \[ \sum_{j=1}^{J} c_{nj} \lambda_j \leq \rho c_{no}, n = 1, \ldots, N; \sum_{j=1}^{J} q_j \lambda_j \leq q_o; \] \[ \sum_{j=1}^{J} y_{mj} \lambda_j \geq y_{mo}, m = 1, \ldots, M; \]

and value-based Russell technical efficiency as

\[ \rho_o^{Russ} = \min_{\rho, s^-, \lambda} 1 - \frac{1}{N} \sum_{n=1}^{N} \frac{s_n^-}{c_{no}} \]

subject to \[ \sum_{j=1}^{J} c_{nj} \lambda_j \leq c_{no} - s_n^-, n = 1, \ldots, N; \sum_{j=1}^{J} q_j \lambda_j \leq q_o; \] \[ \sum_{j=1}^{J} y_{mj} \lambda_j \geq y_{mo}, m = 1, \ldots, M; \]
\[ \sum_{j=1}^{J} \lambda_j = 1; \lambda_j \geq 0, j = 1, \ldots, J; s_n^- \geq 0, n = 1, \ldots, N. \]

Note that if we set \( s_n^- = (1 - \theta_n) c_{no} \), the objective of (12) becomes \( \frac{1}{N} \sum_{n=1}^{N} \theta_n \). Also, if the input cost vector, \( c_{nj} (\forall n, \forall j) \), is replaced by the physical input vector, \( x_{nj} (\forall n, \forall j) \), the objective of (12) becomes an input value-based version of the original Russell measure of technical efficiency defined by Färe and Lovell [8].

When the technical efficiency measures (9), (10), (11), and (12) are feasible and bounded, they are related by the fact that \( \rho^{VTE}_o = \rho^{ERuss}_o \leq \rho^{Russ}_o \leq \rho^{Farr}_o \) for \((c_o, y_o) \in T\). Given these relations we propose two indices of technical efficiency measurement bias. The first is output technical efficiency (OTE) bias, denoted by

\[ B^{OTE}_o = \frac{\rho^{VTE}_o}{\rho^{Russ}_o} \leq 1 \] (13)

for \((c_o, y_o) \in T(q)\). The index of output technical efficiency bias in (13) arises from efficiency measures that ignore output slack. When there is no output slack in the constraints that define the technology for DMU\(_o\), the Russell measure, \( \rho^{Russ}_o \), which accounts for only input slack equals the value-based measure, \( \rho^{VTE}_o \), which accounts for both input cost slack and output slack, and thus \( B^{OTE}_o = 1 \). The second index of technical efficiency bias accounts for the existence of input cost slack (ICS). This bias index is denoted by

\[ B^{ICS}_o = \frac{\rho^{Russ}_o}{\rho^{Farr}_o} \leq 1 \] (14)

for \((c_o, y_o) \in T\) if \( \rho^{Russ}_o \) and \( \rho^{Farr}_o \) are feasible and bounded. The efficient subset of the input cost set is represented by \( \text{Eff}(V; y; q) = \{ c : c' \leq c, c' \neq c, c' \notin V(y; q) \} \). That is, if spending on one input can be reduced and the resulting input cost vector can still feasibly produce a given output, then the original input cost vector is not part of the efficient subset. Input cost slack bias (14) gives the overestimate of value-based technical efficiency from ignoring input cost slack after proportionally contracting all input costs to the input cost isoquant, \( V(c; y) \), but not the cost efficient subset, \( \text{Eff}(V; y; q) \). Value-based technical efficiency (VTE) can be decomposed into the product of technical efficiency bias, input cost slack bias, and value based Farrell technical efficiency:

\[ \rho^{VTE}_o = B^{OTE}_o \times B^{ICS}_o \times \rho^{Farr}_o. \] (15)

We note that \( \gamma^{OSAC} \) is not necessarily less than \( \rho^{VTE}_o \), whereas \( \gamma^{CA} \) must be less than or equal to \( \rho^{Farr}_o \) if each measure is feasible and bounded. Clearly, a DMU is efficient if and only if each efficiency measure equals one. Appendix A1 gives a summary of the cost and technical efficiency measures.

2.3. Fractional programming to linear programming

Since (5) is a fractional programming problem, we follow the Charnes and Cooper [3] procedure to transform a nonlinear programming problem into a linear programming problem by introducing the variable \( t = (1 + \frac{1}{M} \sum_{m=1}^{M} s_m^-)^{-1} \), which is positive for \((x_o, y_o) \in T(q_o)\).
Using \( t \), we convert the decision variables in (2) into new variables as

\[
\begin{align*}
\hat{c}_n &= t c_n, n = 1, \ldots, N, \\
\hat{s}_m^+ &= t s_m^+, m = 1, \ldots, M, \\
\hat{\phi}_m &= t \phi_m, m = 1, \ldots, M, \\
\hat{\lambda}_j &= t \lambda_j, j = 1, \ldots, J.
\end{align*}
\] (16)

Since multiplying both the numerator and denominator in the objective of (5) by the positive scalar \( t \) does not change its value, we rewrite (5) as the following linear programming envelopment form:

\[
\begin{align*}
\rho_o^{OSAC} &= \text{minimize} \quad t, \hat{c}, \hat{s}^+, \hat{\lambda}, \\
& \quad \text{subject to} \\
& \quad \quad t + \frac{1}{M} \sum_{m=1}^{M} \hat{s}_m^+ y_{mo} = 1; \sum_{j=1}^{J} c_{nj} \hat{\lambda}_j \leq \hat{c}_n, n = 1, \ldots, N; \\
& \quad \quad \sum_{j=1}^{J} q_j \hat{\lambda}_j \leq t q_o; \sum_{j=1}^{J} y_{mj} \hat{\lambda}_j = t y_{mo} + \hat{s}_m^+, m = 1, \ldots, M; \\
& \quad \quad \sum_{j=1}^{J} \hat{\lambda}_j = t; \hat{\lambda}_j \geq 0, j = 1, \ldots, J; \\
& \quad \quad \hat{s}_m^+ \geq 0, m = 1, \ldots, M; \hat{c}_n \geq 0, n = 1, \ldots, N; t > 0.
\end{align*}
\] (17)

Obtaining the dual form of (17), we estimate the output slacks-adjusted cost efficiency via the following multiplier form:

\[
\begin{align*}
\text{maximize} \quad & \gamma \\
& \quad \text{subject to} \\
& \quad \quad \gamma + \delta q_o - \sum_{m=1}^{M} \mu_m y_{mo} - \omega \leq 0; \\
& \quad \quad -\sum_{n=1}^{N} \nu_n c_{nj} - \delta q_j + \sum_{m=1}^{M} \mu_m y_{mj} + \omega \leq 0, j = 1, \ldots, J; \\
& \quad \quad \gamma - \mu_m M y_{mo} \leq 0, m = 1, \ldots, M; \nu_n \leq 1/\alpha_{no}, n = 1, \ldots, N; \\
& \quad \quad \nu_n \geq 0, n = 1, \ldots, N; \delta \geq 0; \mu_m \geq 0, m = 1, \ldots, M; \gamma, \omega \text{ free.}
\end{align*}
\] (18)

Note that technology exhibits increasing (decreasing) returns to scale if and only if \( \omega^* > (\prec)0 \) for all optimum solutions where the \( \omega^* \) is associated with a firm on the efficient frontier. Constant returns to scale exist if and only if \( \omega^* = 0 \).

For value-based technical efficiency given by (9) we perform the same data transformation and obtain

\[
\begin{align*}
\rho_o^{VTE} &= \text{minimize} \quad t - \frac{1}{N} \sum_{n=1}^{N} \hat{s}_n^- x_{no} \\
& \quad \text{subject to} \\
& \quad \quad t + \frac{1}{M} \sum_{m=1}^{M} \hat{s}_m^+ y_{mo} = 1; \\
& \quad \quad \sum_{j=1}^{J} c_{nj} \hat{\lambda}_j = t c_{no} - \hat{s}_n^-, n = 1, \ldots, N; \sum_{j=1}^{J} q_j \hat{\lambda}_j = t q_o.
\end{align*}
\] (19)
\[ \sum_{j=1}^{J} y_{mj} \hat{\lambda}_j = ty_{mo} + \hat{s}_m^+, m = 1, \ldots, M; \hat{\lambda}_j \geq 0, j = 1, \ldots, J; \]
\[ \sum_{j=1}^{J} \hat{\lambda}_j = t; \hat{s}_n^- \geq 0, n = 1, \ldots, N; \hat{s}_m^+ \geq 0, m = 1, \ldots, M; t > 0. \]

For the value-based technical efficiency measure given by (9) we perform the same data transformation and estimate the following multiplier form:

\[
\begin{align*}
\text{maximize} & \quad \rho, \nu, \delta, \mu, \omega \\
\text{subject to} & \quad \rho + \sum_{n=1}^{N} \nu_n x_{no} + \delta q_o - \sum_{m=1}^{M} \mu_m y_{mo} - \omega \leq 1; \\
& \quad - \sum_{n=1}^{N} \nu_n c_{nj} - \delta q_j + \sum_{m=1}^{M} \mu_m y_{mj} + \omega \leq 0, j = 1, \ldots, J; \\
& \quad \rho/M y_{mo} - \mu_m \leq 0, m = 1, \ldots, M; \nu_n \geq 1/N c_{no}, n = 1, \ldots, N; \\
& \quad \nu_n \geq 0, n = 1, \ldots, N; \delta \geq 0; \mu_m \geq 0, m = 1, \ldots, M; \rho, \omega \text{ free}. \\
\end{align*}
\]

The problems that estimate \([\gamma_o]^\text{OSAC}, \gamma_o^\text{CA}, \rho_o^\text{Farr}, \rho_o^\text{Russ} \text{ and } \rho_o^\text{VTE}\] are given by equations (18), (4), (11), (12) and (20). The problems that estimate \([\gamma_o^\text{Farr}}, \rho_o^\text{Farr}, \text{ and } \rho_o^\text{Russ}\] are estimated in their primal forms as given by equations (4), (11), and (12).

In our empirical example we consider firms that use only two variable inputs: labor and capital. We are interested in comparing the ratio of labor costs to capital costs that actually occur within a DMU and the ratio of labor costs to capital costs that corresponds with cost efficiency. If multiple solutions to the linear programming problems exist, there may be more than one ratio of labor to capital costs that minimize total costs. To study the relative amount of overspending on labor we define a labor-capital cost mix index as

\[\text{mix} = \frac{\text{labor cost}}{\text{capital cost}} = \frac{c_1^\ast}{c_2^\ast}. \quad \text{(21)}\]

To calculate the relative mix of labor and capital we compute two more LP problems. One is the following maximization program:

\[
\begin{align*}
\text{maximize} & \quad \hat{c}_2 - \hat{c}_1 \\
\text{subject to} & \quad \sum_{i=1}^{N} \hat{c}_n / \alpha_o = \gamma_o^\text{OSAC}; t + \frac{1}{M} \sum_{m=1}^{M} \hat{s}_m^+ = 1; \\
& \quad \sum_{j=1}^{J} c_{nj} \hat{\lambda}_j \leq \hat{c}_n, n = 1, \ldots, N; \sum_{j=1}^{J} q_j \lambda_j \leq tq_o; \\
& \quad \sum_{j=1}^{J} y_{mj} \hat{\lambda}_j = ty_{mo} + \hat{s}_m^+, m = 1, \ldots, M; \hat{\lambda}_j \geq 0, j = 1, \ldots, J; \\
& \quad \sum_{j=1}^{J} \hat{\lambda}_j = t; \hat{s}_n^- \geq 0, m = 1, \ldots, M; \hat{c}_n \geq 0, n = 1, \ldots, N; t > 0.
\end{align*}
\]

where \([\gamma_o]^\text{OSAC}\] is the optimal objective value in (17). Since problem (22) gives the maximal difference between two nonnegative variables, \(\hat{c}_2\) and \(\hat{c}_1\), we can construct the optimal labor-capital cost mix index as \(\text{mix} = \frac{c_1^*}{c_2^*} = \frac{c_1^{1/t^*}}{c_2^{1/t^*}} = \frac{c_1}{c_2}\) where the star (•) are the solutions to (22)
and give a lower bound estimate for the mix index. We obtain an upper bound to the mix index by replacing ‘maximize’ with ‘minimize’ in (22). That is, the upper bound is computed by

\[
\begin{align*}
\text{minimize} & \quad \hat{c}_2 - \hat{c}_1 \\
\text{subject to} & \quad \sum_{i=1}^{N} \frac{\hat{c}_n}{\alpha_o} = \gamma_{OSAC}^o; \quad t + \frac{1}{M} \sum_{m=1}^{M} \hat{s}_m^+ = 1; \\
& \quad \sum_{j=1}^{J} c_{nj}\hat{\lambda}_j \leq \hat{c}_n, n = 1, \ldots, N; \quad \sum_{j=1}^{J} q_j\hat{\lambda}_j \leq tq_o; \\
& \quad \sum_{j=1}^{J} y_{mj}\hat{\lambda}_j = ty_{mo} + \hat{s}_m^+, m = 1, \ldots, M; \quad \hat{\lambda}_j \geq 0, j = 1, \ldots, J; \\
& \quad \sum_{j=1}^{J} \hat{\lambda}_j = t; \quad \hat{s}_m^+ \geq 0, m = 1, \ldots, M; \quad \hat{c}_n \geq 0, n = 1, \ldots, N; \quad t > 0.
\end{align*}
\]

3. An Empirical Illustration: Japanese Securities Firms

We examine the efficiency of Japanese securities firms during fiscal years 2004 to 2006. In the US, financial assets are increasingly held as securities, rather than bank deposits. Hoshi and Kashyap [17] predict that this trend will spill over to Japan and that the Japanese banking industry will decline relative to the securities industry. Given this forecasted increase in the relative importance of the securities industry, a study investigating the efficiency of securities firms is warranted.

Today, Japanese securities firms are the dominant form of non-depository institution performing investment banking in Japan. Article 65 of the Financial System Reform Act prohibits banks from the securities business and specifies four allowable activities for securities firms with a separate license required for each activity: dealing on account, brokerage services, underwriting services, and retail distribution services. Beginning in 1992, the Act allowed banks to form subsidiaries to enter the securities business.

While many researchers have examined the efficiency of Japanese banks, only a small number of researchers have examined the efficiency of Japanese securities firms. Fukuyama and Weber [12] employ DEA methods to examine technical efficiency, allocative efficiency, and cost efficiency of firms in the Japanese securities industry for the 1988-1993 period. They found that the Big Four securities firms (Nomura, Daiwa, Nikko, and Yamaichi) are more cost efficient than smaller securities firms. They also found that non-Big Four securities firms with keiretsu links to banks are more cost efficient than non-Big Four securities firms with keiretsu links to the Big Four securities firms. Hariyama and Okuyama [16] examined the cost structure of Japanese securities industry in order to understand the effects of deregulation on three major securities firms (Nomura, Daiwa, and Nikko) and online securities firms for the period between 1998 and 2002. They found that 50% of online securities firms had significant product-specific economies of scale to exploit for brokerage commissions, although the three major online securities firms did not have further economies of scale to exploit.


\begin{footnote}{One of the big four called Yamaichi Securities opted for voluntary liquidation in November 1997.}

\end{footnote}
choose the amount to earn on each output and pay for each input. The average lost profits
due to inefficiency ranged from 0% to 15% of total assets during the period 1989-2005.

Our data are from the Nikkei Electronic Database System via Financial Quest for the
fiscal years 2004, 2005, and 2006, which end in March of each subsequent year. We confine
our analysis to this period to control for the many changes in the competitive, regulatory,
and technological structure of the Japanese securities industry. We assume that securities
firms produce two outputs using two variables inputs and one non-controllable input. The
first output, $y_1$, is associated with the firm’s brokerage business activities and equals the
sum of the yen value of stock, margin, and bond transactions. The second output, $y_2$, is
associated with the firm’s other business activities that arise from underwriting securities
offerings and handling subscriptions. This second output equals the sum of the yen value of
underwritings of stocks, bonds and certificates plus the yen value of subscriptions of stocks,
bonds, and certificates.

Other researchers have used revenue values as alternative output measures. For example,
Goldberg et al. [15] and Zhang, Zhang, and Luo [28] defined outputs as revenue values for
their studies of US securities firms, and Fukuyama and Weber [12] used revenue values in
their cost/technical efficiency study of Japanese securities firms. Fukuyama and Weber [14]
estimated Japanese securities firms’ profit efficiency using both physical output quantities
and the amount of output earnings.

The two outputs are produced using labor and capital inputs. The amount spent on
labor equals the firm’s personnel expenses, $c_1$. The amount spent on capital, $c_2$, equals the
firm’s real estate related expenses and other expenses related to fixed capital assets. The
yen values of $q, y_1, y_2, c_1,$ and $c_2$ are deflated by the Japanese GDP deflator for each year.

Recent studies of financial institution efficiency have controlled for the risk-return trade-
off faced by managers and the regulatory constraints imposed by national and international
regulators via the Basel Accords. For financial institutions such as banks, the intermediate
output of deposit services helps to finance various assets such as loans and securities. While
some banks might incur higher costs by carefully monitoring the actions of borrowers in
order to preserve equity capital, other banks might trade a reduction in monitoring costs
for higher risk. Because of this potential tradeoff, researchers such as Devaney and Weber
[6], Färe, Grosskopf, and Weber [10], and Fukuyama and Weber [13] controlled for the risk-
return tradeoff in bank efficiency studies by including equity capital as a non-controllable
(or quasi-fixed) input in the production process. The financial and regulatory environment
for securities firms is somewhat different than for banks. Brokerage and underwriting ac-
tivities usually do not require firms to hold securities for long periods and so equity capital,
rather than assets is a better measure of size. (Saunders and Cornett [22]) To control for
firm size and the risk-return tradeoff we include the deflated value of equity capital as an
uncontrollable input in the constraints that define the technology.

Descriptive statistics are provided in Table 1. The sample includes thirty-nine firms in
2004, thirty-eight firms in 2005, and thirty-four firms in 2006, for a total of 111 observations
with no missing values for the variables. Average assets are more than eight times greater
than equity capital. The average firm uses 650 workers and 5 billion yen in tangible and
intangible assets to produce brokerage and underwriting services. Although the average yen
value of brokerage related services is 46 times greater than the yen value of underwriting
services, the revenues generated by brokerage related services are only 43% of total revenues.

Table 2 lists the descriptive statistics for each efficiency measure and bias index. We
also report the number of firms that are efficient for each measure. During 2004 to 2006,
between four and six firms were efficient ($\gamma_0^{OSAC} = 1$) with actual costs equal to minimum
Table 1: Descriptive statistics for observations

<table>
<thead>
<tr>
<th>Year</th>
<th>asset</th>
<th>equity</th>
<th>$c_1$</th>
<th>$c_2$</th>
<th>$y_1$</th>
<th>$y_2$</th>
<th>rev</th>
<th>cost</th>
<th>profit</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004 (39 firms)</td>
<td>511.36</td>
<td>48.99</td>
<td>7.30</td>
<td>8.33</td>
<td>34799.47</td>
<td>634.70</td>
<td>20.97</td>
<td>15.63</td>
<td>4.64</td>
</tr>
<tr>
<td></td>
<td>1384.54</td>
<td>83.22</td>
<td>10.77</td>
<td>13.74</td>
<td>88945.26</td>
<td>1116.94</td>
<td>29.98</td>
<td>24.28</td>
<td>6.04</td>
</tr>
<tr>
<td></td>
<td>1.15</td>
<td>0.80</td>
<td>0.19</td>
<td>0.19</td>
<td>0.42</td>
<td>0.04</td>
<td>0.59</td>
<td>0.41</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>6470.85</td>
<td>419.63</td>
<td>45.87</td>
<td>61.50</td>
<td>460803.71</td>
<td>5129.62</td>
<td>131.74</td>
<td>103.69</td>
<td>23.94</td>
</tr>
<tr>
<td>2005 (38 firms)</td>
<td>516.63</td>
<td>63.75</td>
<td>8.28</td>
<td>9.78</td>
<td>41123.32</td>
<td>994.73</td>
<td>30.97</td>
<td>18.05</td>
<td>12.06</td>
</tr>
<tr>
<td></td>
<td>1521.29</td>
<td>127.23</td>
<td>13.21</td>
<td>18.09</td>
<td>110139.55</td>
<td>1912.34</td>
<td>48.84</td>
<td>30.77</td>
<td>17.82</td>
</tr>
<tr>
<td></td>
<td>1.31</td>
<td>0.86</td>
<td>0.25</td>
<td>0.23</td>
<td>1.39</td>
<td>0.06</td>
<td>0.73</td>
<td>0.52</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>8835.04</td>
<td>754.04</td>
<td>61.89</td>
<td>95.14</td>
<td>648449.46</td>
<td>8653.71</td>
<td>257.17</td>
<td>157.03</td>
<td>93.36</td>
</tr>
<tr>
<td>2006 (34 firms)</td>
<td>540.73</td>
<td>66.47</td>
<td>8.38</td>
<td>11.68</td>
<td>42631.43</td>
<td>964.24</td>
<td>28.65</td>
<td>20.06</td>
<td>7.27</td>
</tr>
<tr>
<td></td>
<td>1588.23</td>
<td>137.74</td>
<td>15.22</td>
<td>24.61</td>
<td>118480.77</td>
<td>1964.39</td>
<td>56.20</td>
<td>39.36</td>
<td>13.73</td>
</tr>
<tr>
<td></td>
<td>1.33</td>
<td>1.12</td>
<td>0.33</td>
<td>0.40</td>
<td>1.19</td>
<td>0.09</td>
<td>0.13</td>
<td>0.78</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>8587.07</td>
<td>778.37</td>
<td>74.10</td>
<td>132.58</td>
<td>669773.95</td>
<td>8159.45</td>
<td>298.86</td>
<td>206.68</td>
<td>71.75</td>
</tr>
<tr>
<td>Pooled (111 firms)</td>
<td>522.16</td>
<td>59.72</td>
<td>7.97</td>
<td>9.85</td>
<td>39363.37</td>
<td>858.89</td>
<td>26.75</td>
<td>17.82</td>
<td>7.99</td>
</tr>
<tr>
<td></td>
<td>1482.42</td>
<td>116.55</td>
<td>12.98</td>
<td>18.94</td>
<td>105054.17</td>
<td>1687.11</td>
<td>45.60</td>
<td>31.46</td>
<td>13.63</td>
</tr>
<tr>
<td></td>
<td>1.15</td>
<td>0.80</td>
<td>0.19</td>
<td>0.19</td>
<td>0.42</td>
<td>0.04</td>
<td>0.13</td>
<td>0.41</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>8835.04</td>
<td>778.37</td>
<td>74.10</td>
<td>132.58</td>
<td>669773.95</td>
<td>8653.71</td>
<td>298.86</td>
<td>206.68</td>
<td>93.36</td>
</tr>
</tbody>
</table>

Legend: \(cost\)=actual costs in billions of yen, \(y_1\)=yen value of stock, margin, and bond transactions in millions of yen, \(y_2\) is the sum of the yen value of underwritings of stocks, bonds and certificates plus the yen value of subscriptions of stocks, bonds, and certificates in millions of yen, \(equity\)=yen value of equity capital in billions of yen, \(assets\)= yen value of assets in billions of yen, \(Assets\), \(equity\), \(revenue\), \(cost\), \(profit\), \(c_1\), \(c_2\), \(y_1\), \(y_2\) are deflated by the GDP deflator.

For these firms it is impossible to proportionally contract labor and capital costs and maintain the same output. Nor is it possible for these same firms to reallocate labor costs and capital costs to further reduce the total cost of production while maintaining the same level of output. Finally, for these same firms, it is impossible to expand any single output at zero cost.

Value-based Farrell technical efficiency \((\rho_{Farr}^{VTE})\) which accounts for slacks, identify the same number of firms as efficient. In 2004 nine firms are efficient for both of these measures and in 2005 and 2006 eight firms are efficient. For the Farrell efficient firms, \(\rho_{Farr}^{VTE} = 1\), which indicates that it is impossible to proportionally contract spending on labor and spending on capital while maintaining the same output. Nor is it possible for these firms to contract the amount spent on any single input or to expand any single output given that \(\rho_{VTE}^{VTE} = 1\). However, we note that average output slacks-adjusted cost efficiency \((\gamma_{OSAC}^{VTE})\) is less than value-based technical efficiency \((\rho_{VTE}^{VTE})\) in each year. This result implies that after proportionally contracting spending on the two inputs and after reducing costs by eliminating input spending slack, the average firm could further reduce costs by reallocating spending on the two inputs, or, it could reduce the slack in one of the outputs.
The estimate of the output slacks-adjusted cost efficiency, $\gamma_{OSAC}^{o}$, gives the simultaneous minimization in total costs relative to the observed costs and the expansion in the two outputs. To illustrate, consider the estimate of cost allocation efficiency for a hypothetical average firm in 2006 in terms of geometric means in Table 3. The mean estimate is $\gamma_{CA}^{o} = 0.353$ and indicates that costs can be decreased by $(1 - 0.353) \times 100\% = 64.7\%$ by reducing input costs proportionally to the input cost isoquant and then reallocating those inputs to achieve cost minimization. The bias in cost allocation efficiency is $B_{CE}^{o} = 0.243$. Given that $B_{CE}^{o} < \gamma_{CA}^{o}$, the presence of output slack means it is possible to expand one of the outputs at zero cost. The decomposition of output slacks-adjusted cost efficiency, $\gamma_{OSAC}^{o} = 0.086$, is the product of cost allocation efficiency and the bias in cost allocation efficiency. This index summarizes all the gains that can be realized by the average securities firm reducing actual costs via a proportional contraction of spending on the two inputs, a non-radial expansion in output by reducing output slack, and a reallocation of spending on the two inputs to achieve cost minimization.

Table 3: Geometric means of efficiency measures

<table>
<thead>
<tr>
<th>Year</th>
<th>$\gamma_{OSAC}^{o}$</th>
<th>$B_{CE}^{o}$</th>
<th>$\gamma_{CA}^{o}$</th>
<th>$\rho_{VTE}$</th>
<th>$B_{OTE}^{o}$</th>
<th>$\rho_{Russ}^{o}$</th>
<th>$B_{ICS}^{o}$</th>
<th>$\rho_{Farr}^{o}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004</td>
<td>0.098</td>
<td>0.278</td>
<td>0.352</td>
<td>0.118</td>
<td>0.247</td>
<td>0.370</td>
<td>0.771</td>
<td>0.480</td>
</tr>
<tr>
<td>2005</td>
<td>0.079</td>
<td>0.239</td>
<td>0.331</td>
<td>0.116</td>
<td>0.334</td>
<td>0.346</td>
<td>0.785</td>
<td>0.441</td>
</tr>
<tr>
<td>2006</td>
<td>0.086</td>
<td>0.243</td>
<td>0.353</td>
<td>0.135</td>
<td>0.366</td>
<td>0.370</td>
<td>0.802</td>
<td>0.462</td>
</tr>
</tbody>
</table>

1. Geometric means are used in order to see the major source of inefficiency.

Table 4 reports output slacks-adjusted cost efficiency scores for all sample firms. DBrain Securities, SBI E Trade Securities, Toyo Securities, and Mitsubishi UFJ Securities are part of

\[ \gamma_{OSAC}^{o} = B_{CE}^{o} \cdot \gamma_{CA}^{o} \cdot \rho_{VTE} = B_{OTE}^{o} \cdot \rho_{Russ}^{o}, \rho_{Russ}^{o} = B_{ICS}^{o} \cdot \rho_{Farr}^{o} \]
the input cost-based technology frontier in every year. In addition, Jet Securities and Shinko Securities help form the frontier in 2004, and Kaneju Securities is part of the frontier in 2006. The brokerage output of Jet Securities decreased from 1316 billion yen in 2004 to 210 billion yen in 2005, causing a substantial reduction in output slacks-adjusted cost efficiency from 1.000 to 0.001. We deleted Jet Securities in 2006 since Nikkei Needs did not report values for $y_1$ and $y_2$ in that year. There is a large gap between the securities firms that define the frontier and the other inefficient securities firms in each year. For instance, in 2004, six firms have $\gamma^{\text{OSAC}}_0 = 1$ and two other firms-Kabu.com Securities and UFJ Tsubasa Securities-have efficiency scores of 0.68 or higher. For the remaining thirty-one firms, $\gamma^{\text{OSAC}}_0$ averages only 0.10. A similar pattern holds in 2005 and 2006, when all but the top six firms have average efficiency of 0.09 and 0.11. Thus, it appears that the securities industry in Japan has a small group of relatively efficient firms and a larger group of inefficient firms.

What is the primary source of output slacks-adjusted cost inefficiency? Is it a failure to minimize spending on the two inputs? Or, is the primary source of inefficiency due to output slack? Table 5 reports the major sources of inefficiency for securities firms. In 2004, thirty-three out of thirty-nine firms were inefficient with $\gamma^{\text{OSAC}}_0 < 1$. In 2005, thirty-four out of thirty-eight firms were inefficient and in 2006, twenty-nine out of thirty-four firms were inefficient. In 2004 and 2005, about 50% of the inefficient firms had $B^{\text{CE}}_o > \gamma^{\text{CA}}_o$. However, by 2006, $\gamma^{\text{CA}}_o$ was greater than $B^{\text{CE}}_o$ for twenty of the twenty-nine inefficient firms, indicating that for most firms, output slack had become the greater source of cost inefficiency than inefficiency caused by a misallocation of inputs. Regarding the value-based technical efficiency decomposition, no firm had its greatest source of inefficiency from bias caused by slacks in the input cost constraints. That is, zero firms had both $B^{\text{ICS}}_o < B^{\text{OTE}}_o$ and $B^{\text{ICS}}_o < \rho^{\text{Farrell}}_o$. For output technical efficiency bias and value-based Farrell technical efficiency, the results are mixed. In 2004 nine firms were efficient with $\rho^{\text{VTE}}_o = 1$ and thirty firms were inefficient. Nineteen out of the thirty inefficient firms had their greatest source of inefficiency due to output slack rather than inefficiency caused by a failure to proportionally contract spending on labor and capital. That is, nineteen out of thirty inefficient firms had $B^{\text{OTE}}_o < \rho^{\text{Farrell}}_o$. However, in 2005 and 2006, more firms (eighteen out of thirty and seventeen out of twenty-six) had their greatest source of value-based technical inefficiency from a failure to proportionally contract spending on labor and capital, rather than a failure to reduce output slack.

To further investigate whether output slack is a significant source of bias in measured cost efficiency we tested whether $\gamma^{\text{OSAC}}_0 = \gamma^{\text{CA}}_0$ using a battery of nonparametric tests. If the two indexes of cost efficiency are the same, then output slack is not a significant source of bias. The test statistics from the Kruskal-Wallis, Median, Savage, Kolmogorov-Smirnov, and Kuiper tests are reported in Table 6. Each test rejects the null hypothesis that $\gamma^{\text{OSAC}}_0 = \gamma^{\text{CA}}_0$ at a 5% significance level. Thus, for Japanese securities firms, indexes of cost efficiency that ignore output slack are biased upward. We also tested whether the proposed components of value-based technical efficiency provide meaningful information. First, we tested whether $\rho^{\text{Russ}}_o = \rho^{\text{Farrell}}_o$ to see whether value-based Farrell technical efficiency is biased by the existence of input cost slack. The Kruskal-Wallis test, Median test, and Kolmogorov-Smirnov test reject the null at a 10% significance level in each year. However, we are unable to reject the null hypothesis in any year at a 10% significance level for the Savage test and Kuiper test. Therefore, for Japanese securities firms there is no definitive evidence that input cost slack is a significant source of input technical inefficiency. Second, we tested whether $\rho^{\text{Russ}}_o = \rho^{\text{VTE}}_o$ to see whether value-based Russell technical efficiency is biased by the existence of output slack. Each test rejected the null hypothesis at a 5% significance
Table 4: Output slacks-adjusted cost efficiency scores, FY2004-FY2006

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ace</td>
<td>0.107</td>
<td>0.096</td>
<td>0.076</td>
</tr>
<tr>
<td>DBBrain Securities</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Mirai Securities</td>
<td>0.008</td>
<td>0.004</td>
<td>0.001</td>
</tr>
<tr>
<td>Jet Securities</td>
<td>1.000</td>
<td>0.001</td>
<td>—</td>
</tr>
<tr>
<td>Aizawa Securities</td>
<td>0.222</td>
<td>0.203</td>
<td>0.253</td>
</tr>
<tr>
<td>Ichiyoshi Securities</td>
<td>0.089</td>
<td>0.077</td>
<td>0.068</td>
</tr>
<tr>
<td>Iwai Securities</td>
<td>0.002</td>
<td>0.004</td>
<td>0.004</td>
</tr>
<tr>
<td>Utsumiya Securities</td>
<td>0.180</td>
<td>0.162</td>
<td>0.147</td>
</tr>
<tr>
<td>SBI E Trade Securities</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Kazaka Securities</td>
<td>0.156</td>
<td>0.223</td>
<td>—</td>
</tr>
<tr>
<td>Kanju Securities</td>
<td>0.107</td>
<td>0.054</td>
<td>1.000</td>
</tr>
<tr>
<td>Sawada Holdings</td>
<td>0.026</td>
<td>0.014</td>
<td>0.003</td>
</tr>
<tr>
<td>Kyokuto Securities</td>
<td>0.051</td>
<td>0.041</td>
<td>0.054</td>
</tr>
<tr>
<td>Kurokawakotoku</td>
<td>0.046</td>
<td>0.025</td>
<td>0.051</td>
</tr>
<tr>
<td>Chuo Securities</td>
<td>0.172</td>
<td>0.155</td>
<td>0.220</td>
</tr>
<tr>
<td>The Kosei Securities</td>
<td>0.004</td>
<td>0.003</td>
<td>0.001</td>
</tr>
<tr>
<td>IDO Securities</td>
<td>0.043</td>
<td>0.043</td>
<td>0.116</td>
</tr>
<tr>
<td>Kabu.com Securities</td>
<td>0.876</td>
<td>0.833</td>
<td>0.811</td>
</tr>
<tr>
<td>Takagi Securities</td>
<td>0.125</td>
<td>0.116</td>
<td>0.142</td>
</tr>
<tr>
<td>The Tachibana Securities</td>
<td>0.061</td>
<td>0.052</td>
<td>0.057</td>
</tr>
<tr>
<td>SBI Securities</td>
<td>0.069</td>
<td>0.058</td>
<td>0.092</td>
</tr>
<tr>
<td>Tokai Tokyo Securities</td>
<td>0.093</td>
<td>0.099</td>
<td>0.051</td>
</tr>
<tr>
<td>UFJ Tsubasa Securities</td>
<td>0.043</td>
<td>0.043</td>
<td>0.116</td>
</tr>
<tr>
<td>Toyo Securities</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Hinode Securities</td>
<td>0.056</td>
<td>0.044</td>
<td>—</td>
</tr>
<tr>
<td>Maeda Securities</td>
<td>0.056</td>
<td>0.046</td>
<td>0.124</td>
</tr>
<tr>
<td>Matsui Securities</td>
<td>0.003</td>
<td>0.001</td>
<td>0.002</td>
</tr>
<tr>
<td>Invast Securities</td>
<td>0.053</td>
<td>0.037</td>
<td>0.028</td>
</tr>
<tr>
<td>Marusan Securities</td>
<td>0.116</td>
<td>0.116</td>
<td>0.107</td>
</tr>
<tr>
<td>Maruhachi Securities</td>
<td>0.207</td>
<td>0.195</td>
<td>0.258</td>
</tr>
<tr>
<td>Mito Securities</td>
<td>0.125</td>
<td>0.140</td>
<td>0.101</td>
</tr>
<tr>
<td>SMBC Friend Securities</td>
<td>0.187</td>
<td>0.149</td>
<td>—</td>
</tr>
<tr>
<td>Mitsubishi UFJ Securities</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Unimat-Yamamaru Securities</td>
<td>0.008</td>
<td>0.007</td>
<td>0.009</td>
</tr>
<tr>
<td>The Yutaka Securities</td>
<td>0.088</td>
<td>0.093</td>
<td>0.189</td>
</tr>
<tr>
<td>Shinko Securities</td>
<td>1.000</td>
<td>0.361</td>
<td>0.427</td>
</tr>
<tr>
<td>Mizuho Investors Securities</td>
<td>0.196</td>
<td>0.111</td>
<td>0.134</td>
</tr>
<tr>
<td>Cosmo Securities</td>
<td>0.244</td>
<td>0.212</td>
<td>0.188</td>
</tr>
</tbody>
</table>

—: Complete data are not available for estimating efficiency in these years.
Table 5: Major sources of inefficiency

<table>
<thead>
<tr>
<th>Year</th>
<th># of firms</th>
<th># of firms with $B^{CE}_o &lt; \gamma^{CA}_o$</th>
<th># of firms with $\rho^{OTE}_o \leq \rho^{ICS}_o$ and $\rho^{OTA}_o &lt; \rho^{Farr}_o$</th>
<th># of firms with $B^{ICS}_o &lt; B^{OTE}_o$ and $\rho^{Farr}_o$</th>
<th># of firms with $\rho^{Farr}_o &lt; B^{ICS}_o$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004</td>
<td>39</td>
<td>17</td>
<td>16</td>
<td>19</td>
<td>0</td>
</tr>
<tr>
<td>2005</td>
<td>38</td>
<td>16</td>
<td>18</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td>2006</td>
<td>34</td>
<td>9</td>
<td>20</td>
<td>9</td>
<td>0</td>
</tr>
</tbody>
</table>

level in 2004 and 2005. In 2006, all tests except for the Savage test rejected the null at a 1% significance level and the Savage test rejected the null at a 10% significance level. These tests indicate that the presence of output slack in the value-based Russell index of input technical efficiency is a significant source of bias.

Table 6: Nonparametric tests for differences in efficiency

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>$\gamma^{RE}_o = \gamma^{CA}_o$</th>
<th>$\rho^{ICTE}_o = \rho^{Farr}_o$</th>
<th>$\rho^{ICTE}_o = \gamma^{TE}_o$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test/Year</td>
<td>2004</td>
<td>2005</td>
<td>2006</td>
</tr>
<tr>
<td>Kruskal-Wallis</td>
<td>19.20</td>
<td>24.82</td>
<td>14.73</td>
</tr>
<tr>
<td>(p-value for $\chi^2$)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>Median</td>
<td>19.20</td>
<td>25.14</td>
<td>11.36</td>
</tr>
<tr>
<td>(p-value for $\chi^2$)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>Savage</td>
<td>7.35</td>
<td>11.23</td>
<td>6.35</td>
</tr>
<tr>
<td>(p-value for $\chi^2$)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>Kolmogorov-Smirnov</td>
<td>2.49</td>
<td>2.75</td>
<td>2.30</td>
</tr>
<tr>
<td>(p-value for $KS_a$)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>Kuiper</td>
<td>2.49</td>
<td>2.75</td>
<td>2.30</td>
</tr>
<tr>
<td>(p-value for $K_a$)</td>
<td>(0.02)</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
</tbody>
</table>

Williamson [27] introduced the idea of expense preference behavior to describe managers who pursue their own self-interest rather than profit maximization or cost minimization. Such behavior is more likely in markets where managers are shielded from competition by regulatory entry barriers or among public sector or non-profit managers who have objectives other than profits. Instead of profits, managers might instead pursue large budgets or a large number of employees because of the power it gives them. In the past, Japanese financial institutions have been shielded from competition by regulatory entry barriers that keep banks out of the securities business. Thus, we are interested in learning whether there is a pattern to the inability of most securities firms to minimize costs.

To investigate potential over or under-spending on labor relative to capital we computed the optimal labor-capital cost mix described in (19) and estimated by (20). The problem of multiple solutions implies that there may be several vectors of labor costs and capital costs that solve the cost minimization problems. Therefore, we obtained a lower bound and upper bound estimate for the ratio of labor costs to capital costs. For each firm the lower and upper bound estimates were the same. In terms of average values, actual labor costs are greater than actual capital costs. However, as shown in Table 7, the ratio of labor cost
to capital cost which is consistent with output slacks-adjusted cost efficiency ranges from only 0.235 to 0.371.

The last row of Table 7 shows the number of firms with an actual labor cost to capital cost ratio less than one and the number of firms with an optimal ratio of labor costs to capital costs less than one. In each year more than 50% of the firms have an actual ratio greater than one. However, the optimal mix is less than one for all but one firm in each year. The optimal mix of labor costs to capital costs for Toyo Securities equals 1.003 in 2004 and equals 1.016 in 2005. Kaneju Securities has an optimal mix equal to 1.695 in 2006, and is the only firm where more should be spent on labor than capital. Only two firms have an optimal ratio of labor costs to capital costs greater than the actual ratio: Jet Securities and Kabu.com Securities. The optimal ratio is less than the actual ratio for all remaining firms. This evidence is consistent with expense preference behavior among Japanese securities firms.

Table 7: Actual and optimal input cost mix

<table>
<thead>
<tr>
<th></th>
<th>2004 (39 firms)</th>
<th>2005 (38 firms)</th>
<th>2006 (34 firms)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Optimal Cost Mix</td>
<td>Actual Cost Mix</td>
<td>Optimal Cost Mix</td>
</tr>
<tr>
<td>mean</td>
<td>0.243</td>
<td>1.096</td>
<td>0.235</td>
</tr>
<tr>
<td>Std. dev.</td>
<td>0.206</td>
<td>0.470</td>
<td>0.201</td>
</tr>
<tr>
<td>min</td>
<td>0.126</td>
<td>0.126</td>
<td>0.100</td>
</tr>
<tr>
<td>max</td>
<td>1.003</td>
<td>2.184</td>
<td>1.016</td>
</tr>
<tr>
<td># of firms with mix &lt; 1</td>
<td>38</td>
<td>14</td>
<td>37</td>
</tr>
</tbody>
</table>

1. The mix index equals the ratio of labor costs to capital costs.

4. Conclusion

In this paper we extended standard slacks-based efficiency measures by introducing an index of output slacks-adjusted cost efficiency. Output slacks-adjusted cost efficiency was decomposed into the product of cost allocation efficiency and an index of cost efficiency bias due to the existence of slack in the output constraints defining the DEA technology. We estimated each efficiency measure using data on Japanese securities firms operating in the period 2004 to 2006. When output slack is ignored, cost efficiency ranges from 33.1% to 35.3%, but when output slacks are included in the model, cost efficiency ranges from only 7.9% to 9.8%. Using nonparametric tests of equality between the components of output slacks-adjusted cost efficiency we found a significant bias and overestimate of cost efficiency from DEA models that ignore output slack.

We also developed and estimated a value-based technical efficiency index. This index was shown to equal the product of an index of output technical efficiency bias, an index of input cost slack bias, and a radial value-based input technical efficiency index. We found that ignoring output slack causes a significant overestimate of overall value-based technical efficiency. However, we found no significant difference in measured efficiency for a radial and a non-radial measure of value-based input technical efficiency.

Finally, we attempted to shed some light on why Japanese securities firms are inefficient. In 1963, Oliver Williamson [27] proposed that the managers of not-for profit firms or firms that are highly regulated might pursue their own self-interest rather than attempt to maximize profits or minimize costs. Such behavior might lead managers to have a preference for
spending on labor, which gives managers greater power, than spending on other inputs. In the past, Japanese securities firms have been shielded from competition by regulations that restrict entry by banks into the securities business. We found clear evidence that Japanese securities firms’ managers have a preference for expenditures on labor, rather than capital. Thus, at least part of the observed cost inefficiency is due to such an expense preference.

### Appendix A1. Definitions of measures

<table>
<thead>
<tr>
<th>Measures and Indexes</th>
<th>Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cost Efficiency</strong></td>
<td><strong>FarrC (Farrell Cost Efficiency)</strong> ( \gamma_{o \text{ Farr}}^\text{Farr} = \min_{\alpha} \left{ \frac{wx}{\alpha} : (x_o, y_o) \in T \right} )</td>
</tr>
<tr>
<td></td>
<td><strong>OSAC (Output Slacks-Adjusted Cost) efficiency</strong> ( \gamma_{o \text{ OSAC}} = \min \left{ \frac{\sum c_n}{\sum \frac{s^+<em>m}{y</em>{mo}}}, (c, y_o + s^+) \in T(q) \right} )</td>
</tr>
<tr>
<td><strong>Cost Efficiency (Measurement) Bias</strong></td>
<td><strong>B_{o \text{ CE}} = \gamma_{o \text{ OSBC}} / \gamma_{o \text{ CA}}</strong></td>
</tr>
<tr>
<td></td>
<td><strong>CA (cost allocation) Efficiency</strong> ( \gamma_{o \text{ CA}} = \min \left{ \sum c_n : (c, y_o) \in T(q) \right} )</td>
</tr>
<tr>
<td><strong>Value-based Technical Efficiency</strong></td>
<td><strong>Value-based Technical Efficiency</strong> ( \gamma_{o \text{ VTE}} = \min_{s^-} \left{ \frac{\sum s^-<em>n / c</em>{no}}{\sum \frac{s^+<em>m}{y</em>{mo}}}, (c_o - s^-, y_o + s^+) \in T(q) \right} )</td>
</tr>
<tr>
<td></td>
<td><strong>(Value-based) Output Technical Efficiency (Measurement) Bias</strong> ( B_{o \text{ OTE}} = \rho_{o \text{ VTE}} / \rho_{o \text{ Russ}} )</td>
</tr>
<tr>
<td></td>
<td><strong>(Value-based) input Cost Slack (Measurement) Bias</strong> ( B_{o \text{ ICS}} = \rho_{o \text{ Russ}} / \rho_{o \text{ Farr}} )</td>
</tr>
<tr>
<td><strong>Additional Value-based Efficiency</strong></td>
<td><strong>Value-based Farrell Technical Efficiency</strong> ( \rho_{o \text{ Farr}}^{\text{Farr}} = \min_{\theta} \left{ \theta : (\theta c_o, y_o) \in T(q) \right} )</td>
</tr>
<tr>
<td></td>
<td><strong>Value-based Russell Technical Efficiency</strong> ( \rho_{o \text{ Russ}}^{\text{Russ}} = \min_{s^-} \left{ \sum s^-<em>n / c</em>{no} : (c_o - s^-, y_o) \in T(q) \right} )</td>
</tr>
</tbody>
</table>

\( \alpha_o \): total cost for DMU_0

### Acknowledgements

An earlier version of this manuscript is presented at the 2008 DEA symposium held at Seikei University, Tokyo. We are grateful to Rolf Färe, Hiroshi Morita and two anonymous referees for their helpful comments. This research is partially supported by the Grants-in-aid for scientific research, fundamental research (A) 18201030 and (B) 19310098, the Japan Society for the Promotion of Science.

### References


Hirofumi Fukuyama  
Faculty of Commerce, Fukuoka University  
8-19-1 Nanakuma, Jonan, Fukuoka 814-0180, Japan  
E-mail: fukuyama@fukuoka-u.ac.jp